
Volume 98
Number 3
2011

Annals
of the
Missouri
Botanical
Garden



WHY ARE WE STILL PRODUCING PAPER FLORAS?¹ *Anthony R. Brach² and David E. Boufford³*

ABSTRACT

Advances in online resources and electronic publication provide the sciences with tools to revolutionize education and research (e.g., cataloging, data archiving and access, and identification). Older journals and monographs are being scanned and increasingly posted online quickly by book scanning projects, although even for new issues, there can be a “moving wall” or lag time (e.g., BioOne and JSTOR). Some scientific disciplines are providing the public with pre-print access to articles in digital libraries, while manuscript availability in other disciplines, such as nomenclatural botany, is often delayed until the time of print publication. Within botany, taxonomic treatments should be provided online at all stages of preparation and revision (in certain cases following initial editing). Now, many published floras can be browsed and searched online (e.g., Flora of North America, Jepson Manual, Flora of Australia). The Flora of China Project provides treatments at all stages, online, prior to printing of volumes. Because of this visibility, the Flora of China Web site regularly receives questions from the general public and helpful review comments from botanists worldwide. With pressing issues of cataloging biodiversity, conservation, and sustainable use of resources, botanists are challenged to prepare and revise online treatments, including interactive identification keys and images, for a worldwide audience of students and researchers.

Key words: digital floras, digital libraries, electronic publication, Flora of China, interactive identification keys, species-id.net.

With the pressing issues of conservation and naming and describing the world’s species (Wheeler et al., 2004), electronic publishing (Lancaster, 1995) provides the tools to revolutionize education and research (e.g., cataloging, data archiving and access, and identification; also see Rosatti & Duncan, 1995). Today, older books, journals, and monographs are being scanned and posted quickly by Google Books

¹ We are thankful to numerous individuals for their discussions of this topic. We thank Flora of China Project contributors, editors, and reviewers for allowing their draft and revised treatments to be posted online. We are grateful to Hong Cui (University of Arizona) for using the Flora of China treatments as a test of the MARTT system and for providing the Greenstone Digital Library, Gregor Hagedorn (Julius Kühn-Institute, Germany) for creating the open data repository on species-id.net, Hong Song (St. Louis University) for incorporating floristic volumes into eFloras.org, Myriam Fica (MO) for assistance in updating the data, and Anne Marie Countie (Harvard University) and Chris Freeland (MO) for their support of servers at the Harvard University Herbaria and Missouri Botanical Garden, respectively. We are thankful for helpful comments by the editor Victoria Hollowell (MO) and three reviewers: Tom Rosatti (JEPS), Peter F. Stevens (MO), and Mike Dallwitz (Giralang, Australia). Support of the U.S. National Science Foundation (BSR-8906215, BSR-9201378, DEB 9626806, DEB 0072682, DEB 0343439, DEB-0514279, DEB-0722217, DEB-0810237, and DEB-0813935) for the Flora of China Project is gratefully acknowledged.

² Missouri Botanical Garden c/o Harvard University Herbaria, 22 Divinity Avenue, Cambridge, Massachusetts 02138-2020, U.S.A. Author for correspondence: brach@oeb.harvard.edu.

³ Harvard University Herbaria, 22 Divinity Avenue, Cambridge, Massachusetts 02138-2020, U.S.A.
doi: 10.3417/2010035

(<<http://books.google.com>>), the Biodiversity Heritage Library (<<http://www.biodiversitylibrary.org>>), and the Botanicus Digital Library (<<http://www.botanicus.org>>). These projects are constructing extensive digital libraries (Greenstein & Thorin, 2002). Even for new issues, though, there can be a “moving wall” or lag time (e.g., BioOne: <<http://www.bioone.org/>> and JSTOR: <<http://www.jstor.org/>>).

Now, several scientific disciplines (e.g., medicine, astronomy, chemistry, and physics) are providing the public with pre-print access to articles (Tomaiuolo & Packer, 2000). Online, peer-reviewed, open-access electronic publication of journal articles has begun in the biological sciences (<<http://www.doaj.org/>>), even sharing archived data supplements to papers, which are not printed because of cost and space issues (Vision, 2010). For the most part, however, botany has lagged behind, despite current examples of online publications (*Constancea*, <<http://ucjeps.berkeley.edu/constancea/>>, The Jepson Manual, <http://ucjeps.berkeley.edu/jepson_flora_project.html>, *Phytotaxa*, <<http://www.mapress.com/phytotaxa>>, *Phytoneuron*, <<http://www.phytoneuron.net/>>, and *Vulpia*, <<http://www.cals.ncsu.edu/plantbiology/ncsc/vulpia>>), and even earlier attempts at computerization (Krauss, 1973). Currently, the issue of valid and effective paperless publication of botanical names is under consideration, predicated by an International Association of Plant Taxonomy’s (IAPT) special committee report (Chapman et al., 2010).

Even earlier, before pre-print status (or post-print, see Harnad, 2003a) of a manuscript, the internet enabled the shift from the trade model (i.e., traditional, restricted commercial publishing) to an online collaborative model for peer review (Harnad, 1996, 2003b; Wilson, 2001). Editors of botanical literature routinely seek scientific reviews by specialists familiar with the concerned plant family and geographic region. To utilize a worldwide audience of reviewers and to connect authors and readers (Tenopir, 1995), taxonomic treatments should be provided online at all stages of preparation and revision (in certain cases after initial editing, e.g., international floras where the author’s native language differs from that of the flora).

Digital libraries provide an answer to overuse of paper and declining funding and bookshelf space. Yet for those who still wish to own a personal, printed copy of a particular book, print on demand (POD) is available (Black, 2011) and is becoming an economical alternative (e.g., <<http://www.printondemand.com>>); however, these are not always of the same print quality and POD can contractually confound issues of

copyright and intellectual property, though these problems are becoming less common. If the copyright issue of providing digital access could be solved for all new publications, it would greatly help botanists.

Electronic paper (Heikenfeld et al., 2011; e.g., <http://en.wikipedia.org/wiki/Amazon_Kindle>) is increasingly popular. Tablet platforms have exploded within the past year, e.g., Motorola’s Xoom with Android’s 3.0 Honeycomb competitive with Apple’s iPad. Botanists have not yet tapped the potential for floristic apps for smart phones and tablets as field tools.

Conventional floras are not designed for use in the field, yet many botanists want to be able to identify plants in the field. Imagine the possibilities of having a flora on a smart phone or tablet and not needing internet access as required by a search engine. Apps extend the taxonomic reach in the field. If the device can be charged, a taxonomist can manage without the internet as long as the flora is loaded. Additionally though, internet access can provide supplemental information (e.g., electronic keys, digitized photos and illustrations, specimen data, etc.).

However, it is noted that some people still prefer carrying around books in which they can occasionally press a specimen within pertinent pages. Other issues include the readability of electronic display screens in bright daylight, the electricity requirements of devices (at least for charging or batteries), cost to produce computers as compared to paper, and long-term archiving (although the Internet Archive is noted here: <<http://www.archive.org/>>).

DIGITAL FLORAS

When paper publications are digitized, previously unavailable functions, such as full text searching and hyperlinking, become possible (Rosatti & Duncan, 1995; Heidorn, 2004; Kirkup et al., 2005; Kress & Krupnick, 2006). Online floras are now providing multiple-field searches into the data (see Brach & Song, 2006), which allow searching by taxon name or distributional fields, such as elevation, province, and country, as well as searches through free text (<http://www.efloras.org/search_page.aspx?flora_id=2>) and linking treatments to related information, photos and illustrations, maps, databases, checklists, and other floras (Funk, 2006). In addition to full text or “context searching,” digital floras also allow for updating with new revised treatments as they become available instead of having to wait for the next paper flora to be published. Data extractions of florulas for geographical areas of special interest, e.g., at a country, state/

province, or county level, are facilitated by digitization (see Kartesz, 2011).

Botanists are challenged to create interactive identification keys (see Dallwitz et al., 2000 onwards; Brach & Song, 2005, <<http://flora.huh.harvard.edu/china/ActKey/>>; Kuoh & Song, 2005, http://www.efloras.org/key_page.aspx?set_id=10083&flora_id=1050; also e.g., Neotropikey: <<http://www.kew.org/science/tropamerica/neotropikey.htm>>). They can provide images to be utilized (e.g., Tropicos image search, <<http://www.tropicos.org/ImageSearch.aspx>>; Biodiversity of the Hengduan Mountains, <<http://hengduan.huh.harvard.edu/fieldnotes>>) by an online audience of students, researchers, and the general public worldwide.

AN EXAMPLE: THE FLORA OF CHINA

The Flora of China Project (FOC; <<http://www.foc.org>>) provides online treatments for vascular plants at all manuscript stages (<http://flora.huh.harvard.edu/china/mss/alphabetical_families.htm>) prior to paper publication of text volumes. Draft manuscripts are noted as such in the header, uploaded to the Web site as HTML files, and replaced with revised versions until considered final for volume publication. Once a particular volume is sent to the printer, the final manuscripts for the volume are then databased (<<http://www.eFloras.org>>, see Brach & Song, 2006). Because of this dynamic accessibility, we regularly receive e-mail from the general public asking about plant identification and use, as well as from specialists worldwide, suggesting corrections and inquiries for particular taxa. Also, after printed publication, the project's editors and coauthors, and specialists outside of the project have the opportunity to submit corrections, which we have routinely incorporated into both the databased, published treatments accessible via an interface on <<http://www.eFloras.org>> and into the project's database of corrigenda for published volumes. Newly published names not available in the printed version are added to the eFloras database and to the FOC Checklist (<<http://www.tropicos.org/Project/FC>>), also see Brach & Song, 2006). The checklist is maintained as a project within the Web-accessible Tropicos (<<http://www.tropicos.org/>>). Because the database is regularly updated with corrections, flora users access a "dynamic" flora via the eFloras interface (while the family PDFs are currently maintained in "static" form so as to provide the original, published treatment). Questions for the future will include how to manage major revisions of families and genera after "print publication" of volumes, alternative treatments

by authors differing in their taxonomic opinions, and authorship of revisions to the database years later.

The FOC Project provides interactive identification keys (Brach & Song, 2005) in DELTA Intkey and NaviKey formats (<<http://flora.huh.harvard.edu/china/ActKey/>>). Electronic keys are available for large genera (generally 50 or more species, but also for some smaller genera). The DELTA data are archived and accessible (<http://www.species-id.net/wiki/Main_Page>) for updates and improvements. The keys can be modified (characters/states and items/taxa) using the DELTA Editor application, and subsequent revised keys posted online.

DEVELOPMENTS IN DIGITIZING FLORAS

Existing natural language descriptions can be parsed by software such as MARTT (Markuper of Taxonomic Treatments) (Cui, 2008) into XML language (demonstrated at <<http://research.sbs.arizona.edu/~hongc/ResearchDemo/MARTTDemo1.html>>), thereby providing searchable databases online (<<http://research.sbs.arizona.edu/gs/cgi-bin/library>>).

For new and revised editions of floras, digital floras should be created starting at the beginning of any new project. Programs such as DDescription Language for TAXonomy (DELTA, Dallwitz, 1980; Dallwitz et al., 1993 onwards; Brach & Song, 2005) provide tools (cf. the DELTA Editor: <<http://delta-intkey.com/www/delta-ed.htm>>) for inputting characters, character states, measurements, and counts into data matrices to then output as natural language descriptions (in which natural language descriptions read as a natural language and not as machine-generated) and interactive identification keys that can be linked to related images. Open repositories for interactive key datasets (see <http://www.species-id.net/wiki/Main_Page>) may be used as examples in building new character sets and interactive keys and in creating new data conversion (Structure of Descriptive Data, SDD; DELTA, etc.) and input utilities, and also to provide an archive for the data. Also, digital image libraries of taxa can be implemented as part of the identification process (Agarwal et al., 2006; Hearn, 2009; also Wikispecies, <<http://species.wikimedia.org/>> and Wikimedia Commons, <<http://commons.wikimedia.org/>>).

With pressing issues of cataloguing biodiversity, conservation, and sustainable use of resources, botanists are now challenged to routinely prepare and revise online treatments, including interactive identification keys and images for a worldwide audience of students and researchers. As botanists,

we should ask ourselves if it is justifiable to continue to produce printed books on paper while at the same time the world is rapidly digitizing older works, and when it appears that soon an increasing number of new works will be available in digital format. Although advances in book scanning, electronic publication, and digital libraries are providing the tools and resources to accomplish this, one has to ask what continuing role the printed flora plays.

Literature Cited

- Agarwal, G., P. Belhumeur, S. Feiner, D. Jacobs, W. J. Kress, R. Ramamoorthi, N. A. Bourg, N. Dixit, H. Ling, D. Mahajan, R. Russell, S. Shirdhonkar, K. Sunkavalli & S. White. 2006. First steps toward an electronic field guide for plants. *Taxon* 55(3): 597–610.
- Black, S. 2011. The magazine rack: Print-on-demand & magazine collection development <http://www.libraryjournal.com/lj/home/889027-264/the_magazine_rack_print-on-demand_.html.csp>, accessed 23 April 2011.
- Brach, A. R. & H. Song. 2005. ActKey: A web-based interactive identification key program. *Taxon* 54(4): 1041–1046.
- Brach, A. R. & H. Song. 2006. eFloras: New directions for online floras exemplified by the Flora of China Project. *Taxon* 55(1): 188–192.
- Chapman, A. D., N. J. Turland & M. F. Watson. 2010. Report of the special committee on electronic publication. *Taxon* 59(6): 1853–1862.
- Cui, H. 2008. Converting taxonomic descriptions to new digital formats. *Biodivers. Informatics* 5: 20–40.
- Dallwitz, M. J. 1980. A general system for coding taxonomic information. *Taxon* 29(1): 41–46.
- Dallwitz, M. J., T. A. Paine & E. J. Zurcher. 1993 onwards. User's Guide to the DELTA System: A General System for Processing Taxonomic Descriptions, 4th ed. CSIRO Division of Entomology, Canberra. <<http://delta-intkey.com/>>, accessed 27 February 2010.
- Dallwitz, M. J., T. A. Paine & E. J. Zurcher. 2000 onwards. Principles of interactive keys. <<http://delta-intkey.com/www/interactivekeys.htm>>, accessed 23 April 2011.
- Funk, V. 2006. Floras: A model for biodiversity studies or a thing of the past? *Taxon* 55(3): 581–588.
- Greenstein, D. I. & S. E. Thorin. 2002. The Digital Library: A Biography. Digital Library Federation, Council on Library and Information Resources, Washington, D.C. <<http://www.clir.org/PUBS/reports/pub109/pub109.pdf>>, accessed 27 February 2010.
- Harnad, S. 1996. Implementing peer review on the net: Scientific quality control in scholarly electronic journals. Pp. 103–108 in R. Peek & G. Newby (editors), *Scholarly Publication: The Electronic Frontier*. MIT Press, Cambridge.
- Harnad, S. 2003a. Eprints: Electronic Preprints and Postprints in M. Drake (editor), *Encyclopedia of Library and Information Science*. Marcel Dekker, Inc., New York. <<http://users.ecs.soton.ac.uk/harnad/Temp/eprints.htm>>, accessed 27 February 2010.
- Harnad, S. 2003b. Open access to peer-reviewed research through author/institution self-archiving: Maximizing research impact by maximizing online access. *J. Postgraduate Med.* 49: 337–342.
- Hearn, D. J. 2009. Shape analysis for the automated identification of plants from images of leaves. *Taxon* 58(3): 934–954.
- Heidorn, P. B. 2004. Publishing digital floras and faunas. *Bull. Amer. Soc. Inform. Sci. Technol.* 30: 2. <<http://www.asis.org/Bulletin/Dec-03/heidorn2.html>>, accessed 27 February 2010.
- Heikenfeld, J., P. Drzaic, J. S. Yeo & T. Koch. 2011. Review paper: A critical review of the present and future prospects for electronic paper. *J. Soc. Inform. Display* 19(2): 129–156. <<http://dx.doi.org/10.1889/JSID19.2.129>>, accessed 23 April 2011.
- Kartesz, J. T. 2011. The Biota of North America Program (BONAP). North American Plant Atlas <<http://www.bonap.org/>> and <<https://rsgis.crrel.usace.army.mil/apex/f?p=703:2:1865983758462951::NO::>>. Chapel Hill, N.C. (maps generated from Kartesz, J. T. 2010. Floristic Synthesis of North America, Version 1.0. Biota of North America Program [BONAP]) (in press), accessed 23 April 2011.
- Kirkup, D., P. Malcolm, G. Christian & A. Paton. 2005. Towards a digital African flora. *Taxon* 54(2): 457–466.
- Krauss, H. M. 1973. The use of generalized information processing systems in the biological sciences. *Taxon* 22(1): 3–18.
- Kress, W. J. & G. A. Krupnick. 2006. The future of Floras: New frameworks, new technologies, new uses. *Taxon* 55(3): 579–580.
- Kuoh, C.-S. & H. Song. 2005. Interactive key to Taiwan grasses using characters of leaf anatomy—The ActKey approach. *Taiwania* 50: 261–271.
- Lancaster, F. W. 1995. The evolution of electronic publishing. *Library Trends* 43(4): 518–527.
- Rosatti, T. J. & T. Duncan. 1995. Floristic information system for California tracheophytes (FISCT). *Madroño* 42(2): 189–196.
- Tenopir, C. 1995. Authors and readers: The keys to success or failure for electronic publishing. *Library Trends* 43(4): 571–591.
- Tomaiuolo, N. G. & J. G. Packer. 2000. Preprint servers: Pushing the envelope of electronic scholarly publishing. *Searcher* 8: 9. <<http://www.infotoday.com/Searcher/oct00/tomaiuolo&packer.htm>>, accessed 27 February 2010.
- Vision, T. J. 2010. Open data and the social contract of scientific publishing. *BioScience* 60(5): 330–331.
- Wheeler, Q. D., P. H. Raven & E. O. Wilson. 2004. Taxonomy: Impediment or expedient? *Science* 303: 285.
- Wilson, H. D. 2001. Informatics: New media and paths of data flow. *Taxon* 50(2): 381–388.